

**Online Appendix for Sterba, S.K. (2013). Understanding linkages among mixture models.
Multivariate Behavioral Research, 48, 775-815.**

Table of Contents.

- I. Full presentation of parallel-process groups-based trajectory (GBT) model.**
- II. Full presentation of parallel-process LTA.**
- III. Multinomial coefficients used to calculate conditional probabilities for empirical example.**
- IV. *Mplus* software syntax for each model described in manuscript.**
- V. Path diagrams for models discussed in manuscript.**
- VI. Brief overview of latent transition analysis (LTA) with 3 timepoints.**

I. Full presentation of parallel-process GBT model.

From manuscript Equation (43), the parallel-process groups-based trajectory model (GBT) is:

$$f(\mathbf{y}_i, \mathbf{z}_i) = \sum_{q=1}^Q \sum_{k=1}^K p(c_i^z = q) p(c_i^y = k | c_i^z = q) f(\mathbf{y}_i | c_i^y = k) f(\mathbf{z}_i | c_i^z = q)$$

Each term in this expression was defined in the manuscript. It was noted in the manuscript that each term could be expressed using earlier-presented manuscript equations, so equations for each term were not repeated to save space. Here, specific equations for each term are presented for added concreteness.

- Marginal probability of class membership for the z -process:

$$p(c_i^z = q) = \exp(\omega^{(q)}) / \sum_{q=1}^Q \exp(\omega^{(q)}) . \text{ For identification } \omega^{(K)} = 0 .$$

- Conditional probability of class membership for the y -process given class membership for the z -process, here written supposing $K=Q=3$:

$$p(c_i^y = k | c_i^z = q) = \frac{\exp(\alpha_k + \beta_{k1}d_{i1} + \beta_{k2}d_{i2})}{\sum_{g=1}^3 \exp(\alpha_g + \beta_{g1}d_{i1} + \beta_{g2}d_{i2})}$$

$Q-1$ dummy variable predictors are used to represent class membership in the z -process (which consists of Q classes). d_{i1} is 1 if z -process class=1, 0 otherwise. Dummy variable d_{i2} is 1 if z -process class=2, 0 otherwise. The denominator summation is over the K outcome classes in the y -process, using summation index g . For identification, coefficients of the last/reference category=0 (here, $\alpha_3 = \beta_{31} = \beta_{32} = 0$).

- Assuming there are J items repeatedly measured for each process, the joint PDF of the y -process repeated measures within class k is given as a product of univariate normal PDFs

$$f(\mathbf{y}_i | c_i^y = k) = \prod_{j=1}^J f(y_{ij} | c_i^y = k)$$

where the parameters $\mu_j^{(k)}$ (a function of estimated growth coefficients) and $\sigma_j^{2(k)}$ of the univariate normal PDF for the j th y -process item are defined as in Equations (21) and (23) in the manuscript.

- The joint PDF of the z -process repeated measures within class q is also given as a product of univariate normal PDFs

$$f(\mathbf{z}_i | c_i^z = q) = \prod_{j=1}^J f(z_{ij} | c_i^z = q)$$

where the parameters $\mu_j^{(z)}$ (a function of estimated growth coefficients) and $\sigma_j^{2(z)}$ of the univariate normal PDF for the j th z -process item are defined as in Equations (21) and (23) in the manuscript.

II. Full presentation of parallel-process LTA.

From manuscript Equation (45):

$$f(\mathbf{y}_i, \mathbf{z}_i) = \sum_{s=1}^S \sum_{m=1}^M \sum_{q=1}^Q \sum_{k=1}^K p(c_{i1}^z=q) p(c_{i1}^y=k | c_{i1}^z=q) p(c_{i2}^z=s | c_{i1}^y=k; c_{i1}^z=q) p(c_{i2}^y=m | c_{i1}^z=q; c_{i1}^y=k; c_{i2}^z=s) \\ p(\mathbf{y}_{i1} | c_{i1}^y=k) p(\mathbf{y}_{i2} | c_{i2}^y=m) p(\mathbf{z}_{i1} | c_{i1}^z=q) p(\mathbf{z}_{i2} | c_{i2}^z=s)$$

Each term in this expression was defined in the manuscript. It was noted in the manuscript that each term could be expressed using earlier-presented manuscript equations, so equations for each term were not repeated to save space. Here, specific equations for each term are presented for added concreteness.

- Marginal probability of time 1 latent state membership for the z-process:

$$p(c_{i1}^z=q) = \exp(\omega^{(q)}) / \sum_{q=1}^Q \exp(\omega^{(q)}) . \text{ For identification } \omega^{(K)} = 0 .$$

- Conditional probability of time 1 state membership for the y-process given time 1 state membership for the z-process, here written supposing at time 1 $K=Q=3$:

$$p(c_{i1}^y=k | c_{i1}^z=q) = \frac{\exp(\alpha_k + \beta_{k1}d_{i1} + \beta_{k2}d_{i2})}{\sum_{g=1}^3 \exp(\alpha_g + \beta_{g1}d_{i1} + \beta_{g2}d_{i2})}$$

$Q-1$ dummy variable predictors are used to represent time 1 latent state membership in the z-process (which consists of Q time 1 states). d_{i1} is 1 if at time 1, z-process state=1, 0 otherwise. Dummy variable d_{i2} is 1 if at time 1, z-process state=2, 0 otherwise. The denominator summation is over the K outcome states (i.e., time 1 states in the y-process), using summation index g . For identification, coefficients of the last/reference category=0 (here, $\alpha_3 = \beta_{31} = \beta_{32} = 0$).

- Conditional probability of time 2 state membership for the z-process given time 1 state membership for the y-process and z-process, here written supposing at $K=Q=S=3$:

$$p(c_{i2}^z=s | c_{i1}^y=k; c_{i1}^z=q) = \frac{\exp(\alpha'_s + \beta'_{s1}d_{i1} + \beta'_{s2}d_{i2} + \beta'_{s3}d_{i3} + \beta'_{s4}d_{i4})}{\sum_{g=1}^3 \exp(\alpha'_g + \beta'_{g1}d_{i1} + \beta'_{g2}d_{i2} + \beta'_{g3}d_{i3} + \beta'_{g4}d_{i4})}$$

Single primes (') are used to indicate that coefficients in this equation differ from those in the earlier-presented multinomial regression (for $p(c_{i1}^y=k | c_{i1}^z=q)$) in this model. Here, $K-1$ dummy variable predictors are used to represent time 1 latent state membership in the y-process (which consists of K time 1 states). d_{i1} is 1 if at time 1, y-process state=1, 0 otherwise. d_{i2} is 1 if at time 1, y-process state=2, 0 otherwise. Additionally, $Q-1$ dummy variable predictors are used to represent time 1 latent state membership in the z-process (which consists of Q time 1 states). d_{i3} is 1 if at time 1, z-process state=1, 0 otherwise. d_{i4} is 1 if at time 1, z-process state=2, 0 otherwise. The denominator summation is over the S outcome states (i.e. time 2 states in the z-process), using summation index g . For identification coefficients of the last/reference category=0 ($\alpha'_3 = \beta'_{31} = \beta'_{32} = \beta'_{33} = \beta'_{34} = 0$).

- Conditional probability of time 2 state membership for the y-process given time 1 state membership for the y-process and time 1 and 2 state membership for the z-process, here written supposing at $K=Q=M=S=3$:

$$p(c_{i2}^y = m | c_{i1}^z = q, c_{i1}^y = k, c_{i2}^z = s) = \frac{\exp(\alpha_m'' + \beta_{m1}'' d_{i1} + \beta_{m2}'' d_{i2} + \beta_{m3}'' d_{i3} + \beta_{m4}'' d_{i4} + \beta_{m5}'' d_{i5} + \beta_{m6}'' d_{i6})}{\sum_{g=1}^3 \exp(\alpha_g'' + \beta_{g1}'' d_{i1} + \beta_{g2}'' d_{i2} + \beta_{g3}'' d_{i3} + \beta_{g4}'' d_{i4} + \beta_{g5}'' d_{i5} + \beta_{g6}'' d_{i6})}$$

Double primes (' ') are used to indicate that coefficients in this equation differ from those in the earlier-presented multinomial regressions in this model. Here, $Q-1$ dummy variable predictors are used to represent time 1 latent state membership in the z -process (which consists of Q time 1 states). d_{i1} is 1 if at time 1, z -process state=1, 0 otherwise. d_{i2} is 1 if at time 1, z -process state=2, 0 otherwise. $K-1$ dummy variable predictors are used to represent time 1 latent state membership in the y -process (which consists of K time 1 states). d_{i3} is 1 if at time 1, y -process state=1, 0 otherwise. d_{i4} is 1 if at time 1, y -process state=2, 0 otherwise. Additionally, $S-1$ dummy variable predictors are used to represent time 2 latent state membership in the z -process (which consists of S time 2 states). d_{i5} is 1 if at time 2, z -process state=1, 0 otherwise. d_{i6} is 1 if at time 2, z -process state=2, 0 otherwise. The denominator summation is over the M outcome states (i.e. time 2 states in the y -process), using summation index g . For identification, coefficients of the last/reference category=0 (here, $\alpha_3'' = \beta_{31}'' = \beta_{32}'' = \beta_{33}'' = \beta_{34}'' = \beta_{35}'' = \beta_{36}'' = 0$).

- Assuming there are J items measured at each timepoint, for each process: the joint PMF of the y -process item responses at time 1 within state k is given as a product of univariate Bernoulli PMFs

$$p(\mathbf{y}_{i1} | c_{i1}^y = k) = \prod_{j=1}^J p(y_{ij1} | c_{i1}^y = k)$$

where the parameter $\pi_{j1}^{(k)}$ of the univariate Bernoulli for the j th y -process item at time 1 (item endorsement probability) is defined as in Equation (24) in the manuscript.

- The joint PMF of the z -process item responses at time 1 within state q is given as a product of univariate Bernoulli PMFs

$$p(\mathbf{z}_{i1} | c_{i1}^z = q) = \prod_{j=1}^J p(z_{ij1} | c_{i1}^z = q)$$

where the parameter $\pi_{j1}^{(q)}$ of the univariate Bernoulli for the j th z -process item at time 1 is defined as in Equation (24) in the manuscript.

- the joint PMF of the y -process item responses at time 2 within state m is given as a product of univariate Bernoulli PMFs

$$p(\mathbf{y}_{i2} | c_{i2}^y = m) = \prod_{j=1}^J p(y_{ij2} | c_{i2}^y = m)$$

where the parameter $\pi_{j2}^{(m)}$ of the univariate Bernoulli for the j th y -process item at time 2 defined as in Equation (25) in the manuscript.

- The joint PMF of the z -process item responses at time 2 within state s is given as a product of univariate Bernoulli PMFs

$$p(\mathbf{z}_{i2} | c_{i2}^z = s) = \prod_{j=1}^J p(z_{ij2} | c_{i2}^z = s)$$

where the parameter $\pi_{j2}^{(s)}$ of the univariate Bernoulli for the j th z -process item at time 2 defined as in Equation (25) in the manuscript.

III. Multinomial coefficients used to calculate conditional probabilities for empirical example.

For the empirical example--a hybrid parallel-process LTA of oppositionality and depressive symptoms allowing across-process local dependence for the shared temper/irritability symptom--one kind of conditional probability of sequence membership was provided in the manuscript. Other kinds of marginal, conditional, and joint probabilities of potential substantive interest can be calculated (see Appendix A) from the following estimated multinomial intercepts and slopes. Notation adheres to that used in the previous section (II) of this online appendix, with the modification that in this example $K=M=Q=S=2$.

Multinomial coefficient	Estimate (SE)
ω^1	0.247 (0.086)
α_1	-3.538 (0.654)
β_{11}	4.825 (0.667)
α'_1	-3.527 (0.624)
β'_{11}	0.901 (0.473)
β'_{12}	3.981 (0.747)
α''_1	-3.365 (0.697)
β''_{11}	-7.772 (15.893)
β''_{12}	10.859 (15.900)
β''_{13}	10.762 (15.892)

Notes. For identification, coefficients in the last/reference category were zero:

$$\omega^2 = 0; \alpha_2 = \beta_{21} = 0; \alpha'_2 = \beta'_{21} = \beta'_{22} = 0; \alpha''_2 = \beta''_{21} = \beta''_{22} = \beta''_{23} = 0).$$

IV. *Mplus* software syntax for each model described in manuscript.

The following *Mplus* syntax (L.K. Muthén & B.O. Muthén, 1998-2012) is provided for all models described in the manuscript. Although syntax for some of these models is also available in the *Mplus* users guide, that syntax relies more upon defaults for convenience. Here, for pedagogical purposes, all estimated parameters within-class are explicitly represented in the syntax, so the syntax more directly mirrors the equation presented in the manuscript.

Regarding latent state or transition probabilities from LTA, parallel-process LTA, and parallel-process GBT (that were described in the manuscript's Appendix A), some are computed automatically by *Mplus*, from estimated multinomial coefficients. Alternatively or in addition, all of these probabilities can be computed by defining them as a function of estimated multinomial logistic coefficients in a MODEL CONSTRAINT command, using equations from the manuscript. *Mplus* 7.0 allows more of these probabilities to be automatically computed or requested, and also allows referring to the probabilities directly in the input (using parameterization=probability).

Specifics on the below syntax:

- Syntax for each model contains two classes (or two states per timepoint).
- When there are multiple processes, each process has two classes (or two states per timepoint).
- Where applicable, measurement invariance is imposed.
- When residual and/or factor variances are estimated, they are constrained equal across-class.
- For mixture models with one categorical latent variable, the between model (Equation (8)) is estimated automatically. In the UFMN syntax it is referred to explicitly for pedagogical purposes, and then in other models with one categorical latent variable, it is omitted.
- Syntax for outputting posterior probabilities is the same for all models, so it is only shown for the UFMN.
- Note that multinomial coefficients (intercepts or slopes) in the reference (last) class are never referred to in the syntax because they are automatically constrained to 0 for identification purposes, as discussed in the manuscript.
- Plots like manuscript Figures 3-4 can be obtained by adding the PLOT command to the end of the input (e.g., *PLOT: TYPE=PLOT1 PLOT2 PLOT3;*) and listing the outcomes in the SERIES subcommand (e.g. *SERIES IS y1-y5 (*);* for a single process or *SERIES IS y1-y5 (*) / z1-z5 (*);* for two processes).

Univariate finite normal mixture:

```

DATA: FILE IS yourdata.dat;
VARIABLE:
NAMES ARE y;
USEVARIABLES ARE y;
CLASSES = class (2);
!request K=2
ANALYSIS: TYPE=MIXTURE;
ESTIMATOR=ML;
!request maximum likelihood
STARTS = 500 50;
!request random starting values
MODEL:
%overall%
[class#1]; !K-1 multinomial int.
!from Eqn(8), automatically included
y; !variance
%class#1%
[y]; !mean in class 1
%class#2%
[y]; !mean in class 2

SAVEDATA: FILE IS
2class_posteriorprob.dat;
SAVE ARE cprobabilities;
!request posterior probabilities

```

Latent profile analysis:

```

DATA: FILE IS yourdata.dat;
VARIABLE:
NAMES ARE y1-y5;
USEVARIABLES ARE y1-y5;
CLASSES = class (2);
ANALYSIS: TYPE=MIXTURE;
ESTIMATOR=ML;
STARTS = 500 50;

MODEL:
%overall%
y1-y5; !J variances
%class#1%
[y1-y5]; !J means in class 1
%class#2%
[y1-y5]; !J means in class 2

```

Latent class analysis:

```

DATA: FILE IS yourdata.dat;
VARIABLE:
NAMES ARE y1-y4;
USEVARIABLES ARE y1-y4;
CATEGORICAL ARE y1-y4;
CLASSES = class (2);
ANALYSIS: TYPE=MIXTURE;
ESTIMATOR=ML;
STARTS = 500 50;

MODEL:
%overall%
%class#1%
[y1$1-y4$1]; !J thresholds class 1
%class#2%
[y1$1-y4$1]; !J thresholds class 2

```

Groups-based trajectory model:

```

DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y1-y5;
USEVARIABLES ARE y1-y5;
classes = class(2);
ANALYSIS: TYPE = MIXTURE;
ESTIMATOR=ML; STARTS = 500 50;

MODEL:
%overall%
int by y1-y5@1;
!define intercept coefficient, loadings are 1's
linear by y1@0 y2@1 y3@2 y4@3 y5@4;
!define linear coefficient, loadings are time scores
quad by y1@0 y2@1 y3@4 y4@9 y5@16;
!define quadratic coeff., loadings are sq. time scores
[y1-y5@0];
y1-y5 (1) ; !J residual variances, equal across time
int@0; linear@0; quad@0; !no factor variances
int with linear@0; int with quad@0; linear with
quad@0; !no factor covariances
%class#1%
[int linear quad]; !growth coeff means in class 1
%class#2%
[int linear quad]; !growth coeff means in class 2

```

Latent transition analysis:

```

DATA: FILE IS yourdata.dat;
VARIABLE:
NAMES ARE y11 y21 y31 y41 y12 y22 y32 y42;
USEVARIABLES ARE
y11 y21 y31 y41 y12 y22 y32 y42;
CATEGORICAL ARE
y11 y21 y31 y41 y12 y22 y32 y42;
classes = c1 (2) c2 (2);
!request K=2 and M=2
ANALYSIS: TYPE = MIXTURE; ESTIMATOR=ML;
STARTS = 500 50;

MODEL:
%overall%
c2#1 on c1#1;
!multinomial slope, regress t state1 on t-1 state1 (Eq30)
[c2#1]; !multinomial intercept for t state1 (Eq30)
[c1#1]; !multinomial intercept, for t-1 (initial) state 1
model c1:
%c1#1%
[y11$1] (1); [y21$1] (2); [y31$1] (3); [y41$1] (4);
!J thresholds held equal within-state across-time -- MI
%c1#2%
[y11$1] (5); [y21$1] (6); [y31$1] (7); [y41$1] (8);
model c2:
%c2#1%
[y12$1] (1); [y22$1] (2); [y32$1] (3); [y42$1] (4);
%c2#2%
[y12$1] (5); [y22$1] (6); [y32$1] (7); [y42$1] (8);

```

Note that for hybrid mixture models, there are several options for identifying the latent factors, and many options for what parameters to allow to vary across class. Below one option is shown per model.

Factor mixture model:

```
DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y1-y5;
USEVARIABLES ARE y1-y5;
CLASSES = class (2);
ANALYSIS: TYPE= MIXTURE;
ESTIMATOR=ML; STARTS = 500 50;

MODEL:
%overall%
fac1 by y1-y5*; !J factor loadings
[fac1@0]; !factor mean fixed to 0
fac1@1; !factor variance fixed to 1
y1-y5; !J residual variances
%class#1%
[y1-y5]; !J outcome intercepts, class 1
%class#2%
[y1-y5]; !J outcome intercepts, class 2
```

Categorical-item factor mixture model:

```
DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y1-y4;
USEVARIABLES ARE y1-y4;
CATEGORICAL ARE y1-y4;
CLASSES = class (2);
ANALYSIS: TYPE= MIXTURE; ESTIMATOR=ML;
STARTS = 500 50; algorithm=integration;

MODEL:
%overall%
fac1 by y1-y4*; !J factor loadings
[fac1@0]; !factor mean fixed to 0
fac1@1; !factor variance fixed to 1
%class#1%
[y1$1-y4$1]; !J outcome thresholds, class 1
%class#2%
[y1$1-y4$1]; !J outcome thresholds, class 1
```

Growth mixture model:

```
DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y1-y5;
USEVARIABLES ARE y1-y5;
CLASSES = class (2);
ANALYSIS: TYPE= MIXTURE;
ESTIMATOR=ML; STARTS = 500 50;

MODEL:
%overall%
int by y1-y5@1;
!define intercept coefficient, loadings are 1's
linear by y1@0 y2@1 y3@2 y4@3 y5@4;
!define linear coefficient, loadings are time scores
quad by y1@0 y2@1 y3@4 y4@9 y5@16;
!define quadratic coeff., loadings are sq. time scores
[y1-y5@0]; !J outcome intercepts, fixed to 0
y1-y5 (1) ; !J residual variances, equal across time
linear@0; quad@0; !no linear or quad variances
int with linear@0; int with quad@0;
linear with quad@0; !no factor covariances here
%class#1%
[int linear quad]; !growth coeff means in class 1
int; !intercept factor variance, class 1
%class#2%
[int linear quad]; !growth coeff means in class 2
int; !intercept factor variance, class 2
```

Latent transition model with a categorical-item factor measurement model:

```
DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y11 y21 y31 y41 y12 y22
y32 y42; USEVARIABLES ARE y11 y21 y31 y41
y12 y22 y32 y42; CATEGORICAL ARE
y11 y21 y31 y41 y12 y22 y32 y42;
classes = c1 (2) c2 (2);
ANALYSIS: TYPE= MIXTURE; ESTIMATOR=ML;
STARTS = 500 50;

MODEL:
%overall%
fac1 by y11* y21 y31 y41; !J factor loadings, time 1
fac2 by y12* y22 y32 y42; !J factor loadings, time 2
fac1@1; fac2@1; !factor variances fixed to 1
fac1 with fac2@0; !factor cov here fixed to 0
[fac1@0 fac2@0]; !factor means fixed to 0
c2#1 on c1#1;
!multinomial slope, regress t state1 on t-1 state1 (Eq30)
[c2#1]; !multinomial intercept for t state1 (Eq30)
[c1#1]; !multinomial intercept, for t-1 (initial) state1

model c1:
%c1#1%
[y11$1] (1); [y21$1] (2); [y31$1] (3); [y41$1] (4);
!J thresholds held equal within-state across-time -- MI
%c1#2%
[y11$1] (5); [y21$1] (6); [y31$1] (7); [y41$1] (8);

model c2:
%c2#1%
[y12$1] (1); [y22$1] (2); [y32$1] (3); [y42$1] (4);
%c2#2%
[y12$1] (5); [y22$1] (6); [y32$1] (7); [y42$1] (8);
```

Parallel-process latent class growth model:

```

DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y1-y5 z1-z5;
USEVARIABLES ARE y1-y5 z1-z5;
  classes = cy (2) cz (2);
  !request K=2 and Q=2
ANALYSIS: TYPE = MIXTURE;
ESTIMATOR=ML; STARTS = 500 50;
MODEL:
%overall%
cy#1 on cz#1;
!multinomial slope, regress yclass1 on zclass1
[cz#1]; !multinomial intercept for zclass1
[cy#1]; !multinomial intercept for yclass1

!define intercept, linear, quad coeff. for yprocess
inty by y1-y5@1;
lineary by y1@0 y2@1 y3@2 y4@3 y5@4;
quady by y1@0 y2@1 y3@4 y4@9 y5@16;
[y1-y5@0]; !J intercepts, fixed to 0, in y process
y1-y5 (1); !J equal res var across time, in y process

!define intercept, linear, quad coeff. for zprocess
intz by z1-z5@1;
linearz by z1@0 z2@1 z3@2 z4@3 z5@4;
quadz by z1@0 z2@1 z3@4 z4@9 z5@16;
[z1-z5@0]; !J intercepts, fixed to 0, in z process
z1-z5 (2); !J equal res var across time, in z process

inty lineary quady intz linearz quadz with
intz@0 linearz@0 quadz@0 inty@0 lineary@0
quady@0; !no factor covariances
inty@0 lineary@0 quady@0 intz@0 linearz@0
quadz@0; !no factor variances

model cy:
%cy#1%
[inty lineary quady]; !y process growth means class1
%cy#2%
[inty lineary quady]; !y process growth means class2

model cz:
%cz#1%
[intz linearz quadz]; !z process growth means class1
%cz#2%
[intz linearz quadz]; !z process growth means class2

```

Parallel-process latent transition analysis:

```

DATA: FILE IS yourdata.dat;
VARIABLE: NAMES ARE y11 y21 y31 y41 y12 y22
y32 y42 z11 z21 z31 z41 z12 z22 z32 z42;
USEVARIABLES ARE y11 y21 y31 y41 y12 y22
y32 y42 z11 z21 z31 z41 z12 z22 z32 z42;
CATEGORICAL ARE y11 y21 y31 y41 y12 y22 y32
y42 z11 z21 z31 z41 z12 z22 z32 z42;
  classes = c1z (2) c1y (2) c2z (2) c2y (2);
  !request Q=2 and K=2 and S=2 and M=2
ANALYSIS: TYPE = MIXTURE; ESTIMATOR=ML;
STARTS = 500 50;
MODEL:
%overall%
[c2y#1]; !multinomial intercept for t ystate1
[c1y#1]; !multinomial intercept for t-1 ystate1
[c2z#1]; !multinomial intercept for t zstate1
[c1z#1]; !multinomial intercept for t-1 zstate1
!multinomial slopes:
c1y#1 on c1z#1;
! regress t-1 ystate1 on t-1 zstate1
c2z#1 on c1z#1 c1y#1;
! regress t zstate1 on t-1 zstate1, t-1 ystate1
c2y#1 on c1z#1 c1y#1 c2z#1;
! regress t ystate1 on t-1 zstate1, t-1 ystate1, t zstate1

model c1z:
%c1z#1%
[z11$1] (11); [z21$1] (22);
[z31$1] (33); [z41$1] (44);
!J thresholds=within-state within-process across-time: MI
%c1z#2%
[z11$1] (55); [z21$1] (66);
[z31$1] (77); [z41$1] (88);

model c1y:
%c1y#1%
[y11$1] (1); [y21$1] (2);
[y31$1] (3); [y41$1] (4);
%c1y#2%
[y11$1] (5); [y21$1] (6);
[y31$1] (7); [y41$1] (8);

model c2z:
%c2z#1%
[z12$1] (11); [z22$1] (22);
[z32$1] (33); [z42$1] (44);
%c2z#2%
[z12$1] (55); [z22$1] (66);
[z32$1] (77); [z42$1] (88);

model c2y:
%c2y#1%
[y12$1] (1); [y22$1] (2);
[y32$1] (3); [y42$1] (4);
%c2y#2%
[y12$1] (5); [y22$1] (6);
[y32$1] (7); [y42$1] (8);

```

Empirical example analysis model:

Syntax is notated in blue with changes from the above parallel process LTA syntax.

```

DATA: FILE IS empiricalex.dat;
VARIABLE: NAMES ARE disob1 temper1 guilt1 whine1 saddep1 slowmv1 irrit1 conc1
  disob2 temper2 guilt2 whine2 saddep2 slowmv2 irrit2 conc2;
USEVARIABLES ARE disob1 temper1 guilt1 whine1 saddep1 slowmv1 irrit1 conc1
  disob2 temper2 guilt2 whine2 saddep2 slowmv2 irrit2 conc2;
CATEGORICAL ARE disob1 temper1 guilt1 whine1 saddep1 slowmv1 irrit1 conc1
  disob2 temper2 guilt2 whine2 saddep2 slowmv2 irrit2 conc2;
MISSING ARE .;
classes = c1z (2) c1y (2) c2z (2) c2y (2);
ANALYSIS: TYPE= MIXTURE; ESTIMATOR=ML;
STARTS = 500 50; algorithm=integration; !numerical integration is necessary

MODEL:
%overall%
[c2y#1] ;
[c1y#1] ;
[c2z#1] ;
[c1z#1] ;
c1y#1 on c1z#1;
c2z#1 on c1z#1 c1y#1;
c2y#1 on c1z#1 c1y#1 c2z#1;
fac1 by irrit1@1 temper1 ; ! one y-process symptom & 1 z process symptom load on time t-1 factor
! one factor loading is fixed for identification and the other factor loading is estimated
fac2 by irrit2@1 temper2; ! one y-process symptom & 1 z process symptom load on time t factor
! one factor loading is fixed for identification and the other factor loading is estimated
fac1@1; fac2@1; ! factor variances are fixed to 1 for identification
fac1 with fac2@0; ! factor covariance is fixed to 0 (optional)
[fac1@0 fac2@0]; ! factor means are fixed to 0

model c1z:
%c1z#1%
[ disob1$1] (11); [ temper1$1] (22); [ guilt1$1] (33); [ whine1$1] (44);
%c1z#2%
[ disob1$1] (55); [ temper1$1] (66); [ guilt1$1] (77); [ whine1$1] (88);

model c1y:
%c1y#1%
[ saddep1$1] (1); [ slowmv1$1] (2); [ conc1$1] (3); [ irrit1$1] (4);
%c1y#2%
[ saddep1$1] (5); [ slowmv1$1] (6); [ conc1$1] (7); [ irrit1$1] (8);

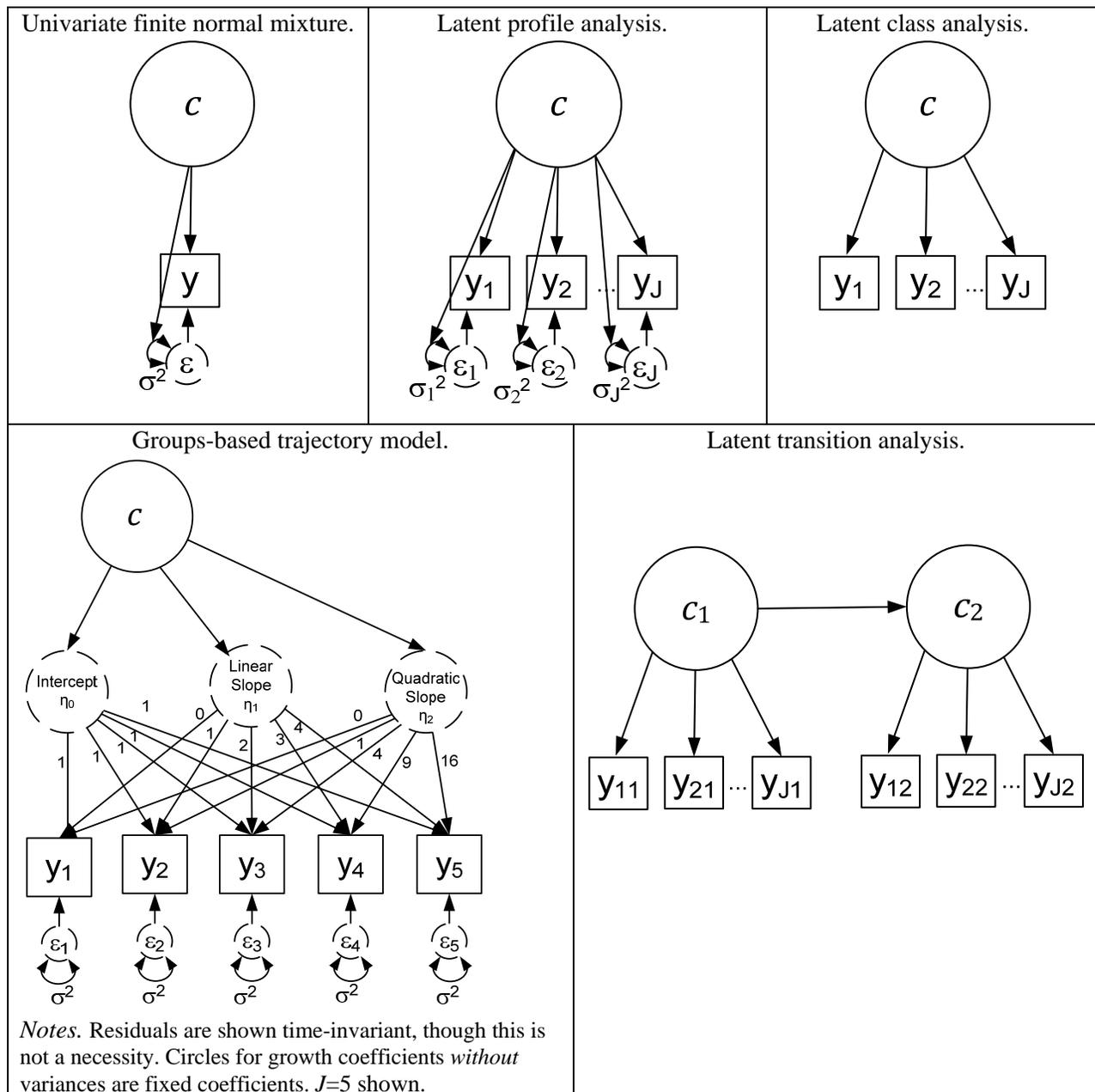
model c2z:
%c2z#1%
[ disob2$1] (11); [ temper2$1] (22); [ guilt2$1] (33); [ whine2$1] (44);
%c2z#2%
[ disob2$1] (55); [ temper2$1] (66); [ guilt2$1] (77); [ whine2$1] (88);

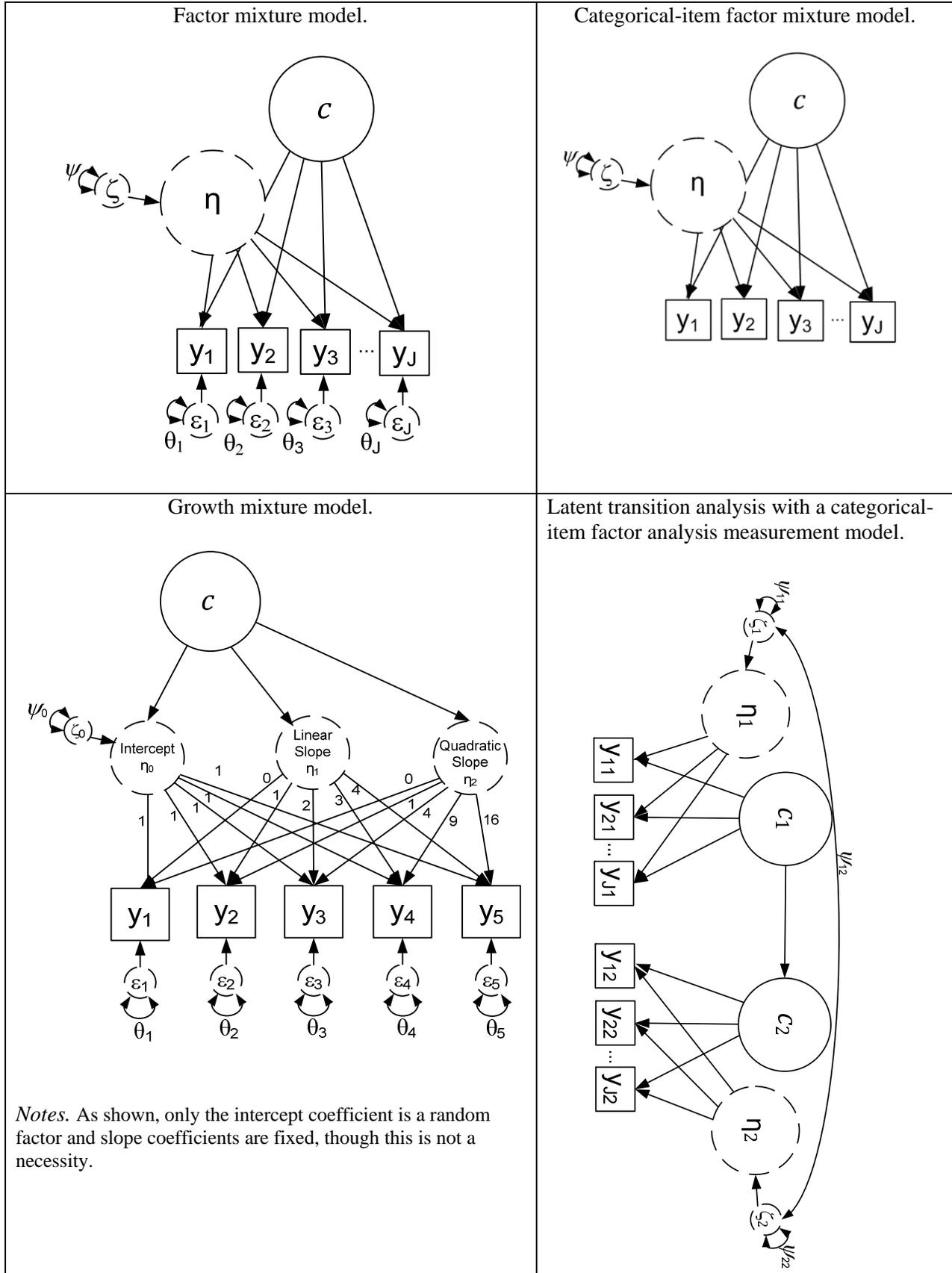
model c2y:
%c2y#1%
[ saddep2$1] (1); [ slowmv2$1] (2); [ conc2$1] (3); [ irrit2$1] (4);
%c2y#2%
[ saddep2$1] (5); [ slowmv2$1] (6); [ conc2$1] (7); [ irrit2$1] (8);

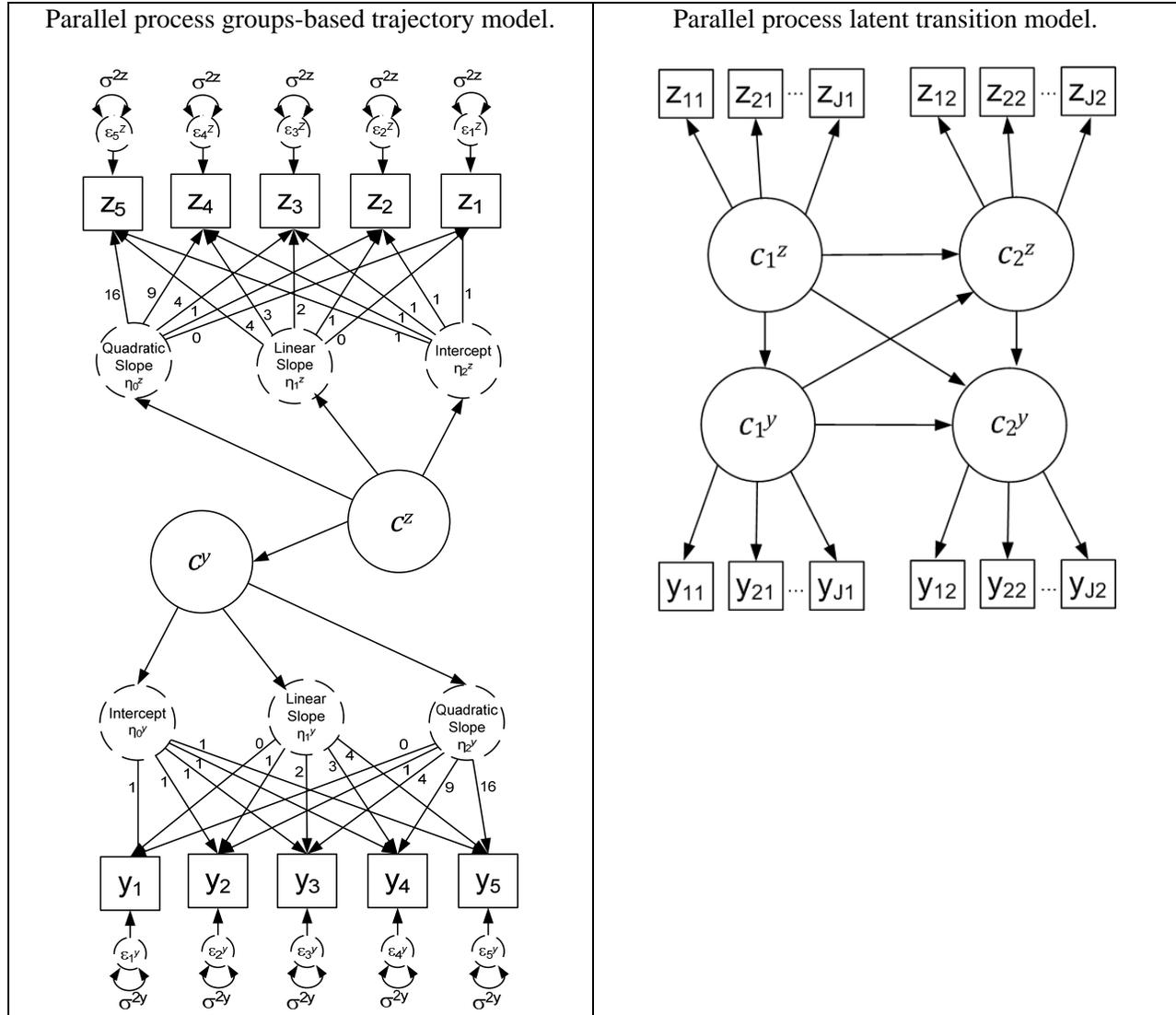
```

V. Path diagrams for models discussed in manuscript.

There is not an established path diagrammatic convention in mixture modeling that allows model equations to be exactly reconstructed from diagrams using, say, Wright's tracing rules, as there is in structural equation modeling. Rather, path diagram-like representations are used in a heuristic way for mixture models to help visualize the model, but without an exact mapping to the model equations. Here, a solid circle denoted with c represents a categorical latent variable; dashed circles represent continuous latent variables. A straight arrows represent a regression path (that may be linear or nonlinear). The categorical latent variable points at parameters that are allowed to be class-varying--with the exception that when it points to a measured variable (square) or a latent factor, the corresponding mean is allowed to vary across class. (For simplicity, item intercepts/thresholds or factor means are not represented.) Curved arrows represent variances or covariances. J refers to the number of outcomes.







VI. Brief overview of LTA with 3 timepoints.

LTA with two timepoints is used in the manuscript. A LTA with three timepoints is briefly summarized here. At time 3 items $y_{i13} - y_{i33}$ serve as indicators of c_{i3} , which has H latent states. \mathbf{y}_{i3} is the response pattern for J items at time 3. Under a first-order Markov assumption, the combined LTA model for $T=3$ would entail adding a within-state model for time 3, similarly to Equation (29), and adding a multinomial regression of latent states at time 3 on latent states at time 2 similarly to Equation (30) (implying an additional transition probability matrix).

$$p(\mathbf{y}_i) = \sum_{h=1}^H \sum_{m=1}^M \sum_{k=1}^K (p(c_{i1} = k) p(c_{i2} = m | c_{i1} = k) p(c_{i3} = h | c_{i2} = m) p(\mathbf{y}_{i1} | c_{i1} = k) p(\mathbf{y}_{i2} | c_{i2} = m) p(\mathbf{y}_{i3} | c_{i3} = h))$$

Probabilities of state membership at time 3 are calculable from estimated multinomial coefficients following principles in Appendix A.

LTAs with $T \geq 3$ allow the examination of stationarity of transition probabilities across time. They also allow the possibility of investigating a second-order Markov process (regressing latent states at time t not just on states at $t-1$ but also on states at $t-2$), although the number of added parameters required may be prohibitive. These possibilities are not demonstrated here.